Tales from the Terminal Room
February 2006, Issue No. 68

Tales from the Terminal Room (TFTTR) is a monthly newsletter, with the exception of July and August , and November and December, which are published as single issues. TFTTR includes reviews and comparisons of information sources; updates to the RBA Web site Business Sources and other useful resources; dealing with technical and access problems on the Net; and news of RBA's training courses and publications.

Tales from the Terminal Room can be delivered via email as plain text or as a PDF with active links. You can join the distribution list by going to http://www.rba.co.uk/tfttr/ and filling in the form. You will be sent an email asking you to confirm that you want to be added to the list. TFTTR is also available as an RSS feed. The URL for the feed is http://www.rba.co.uk/rss/tfttr.xml
.

In this issue:

Search Tools

Jeeves and Teoma Retire

Jeeves's retirement from Ask Jeeves has been well publicised - the search engine now answers to http://www.ask.com/ or http://www.ask.co.uk/ - but many searchers may be unaware that Teoma has also been put out to grass. Try to connect to teoma.com and you are now redirected to Ask.

Ask Jeeves acquired Teoma in late 2001 since when it has been the technology underlying Ask Jeeves. Many of the Teoma features are now in Ask; the only feature that is not is the "Resources", or hubs, pages. Ask say that given that this receives less than a 1% click-rate, they don't think many people will miss it. If you were put off Ask Jeeves in its early days do give it another try. It has improved immensely over the last couple of years with some very useful features such as Zoom, which suggests ways of broadening or narrowing down your search. It also has a very good image search option.

[Originally posted by Karen Blakeman to the UKeiG blog - http://www.ukeig.org.uk/blog/2006/02/jeeves-and-teoma-retire.html]

Google Desktop 3 launched

http://desktop.google.com/

Regular readers if TFTTR are well aware of my concerns regarding earlier versions of this software. Rather than just create an index of the original files on your computer, as the other desktop search tools do, Google desktop makes text copies of your documents and keeps them in a cache on your PC. It is this cache that Google Desktop indexes. The main problem I have with this is that if you delete the original document a copy remains in the desktop cache. There is a remove facility but it is not straightforward and certainly not foolproof. For many organisations, this may contravene their document retention and/or records management policies and has serious implications regarding data protection and FOI for some UK users.

Google Desktop 3 goes a stage further and enables you to search for files across all your computers. Great, you might think. If you are on the road and have forgotten to copy a vital document from your desktop machine onto your laptop you can quickly get to it via Desktop 3. Whoa there! In order to do this Google stores your files (web history, cache, office documents, PDF and text documents) on its own servers. That is something I am sure most organisations would definitely not want to happen.

In the US the Electronic Frontier Foundation has urged consumers to boycott the software, warning that Google could be forced to turn over the data to the government if subpoenaed, even if the data is stored on Google servers temporarily. Google says that your files may remain on its servers for up to 30 days but several commentators have pointed out that Google's desktop search privacy policy states that if you uninstall the Google desktop, or deactivate your Google account, some data may stay on the Google's servers for up to 60 days.

The feature is turned off by default but I noticed that when I opted to index my Googlemail, the same screen had the Search Across Computers option, which automatically ticked itself when I enabled googlemail indexing. OK, so you can untick it but one might be tempted to leave it, especially if you have not read the further information or the privacy policy. At least you must have a Google account of some sort to use this feature and you have to type the details into Google Desktop, so that is a check against accidentally enabling it.

If you are really interested in desktop search, there are plenty of others around. I use Yahoo, but Copernic and Exalead are also excellent, and I have received good reports on them from other users.
[Originally posted by Karen Blakeman to the UKeiG blog http://www.ukeig.org.uk/blog/2006/02/google-desktop-3-launched.html]

Exalead reaches 4 billion pages

http://www.exalead.com/

Exalead is now searching over four billion Web pages and aims to reach eight billion by July. As many information professionals say the size of the web database is irrelevant if it generates rubbish results, but a critical mass is essential to encourage users to use a search engine and Exalead has certainly done that now.

I am a great fan of Exalead: it has several unique features and has resurrected a couple that were discarded by the other mainstream engines long ago. You can use wildcards (*) at the end of a word and the NEAR command looks for your terms within 16 words of each other. Unique features include phonetic and approximate spelling search options, and the pattern matching option enables you to cheat - ahem, I mean - can help you find that final, elusive solution for your crossword puzzle.

And if you want to run your search in other search tools after running it through Exalead, you can set up short cuts to them from Exalead's home page. (For further details see Tales From the Terminal Room October 2005: Exalead revamps web search http://www.rba.co.uk/tfttr/archives/2005/oct2005.shtml)

Blogs & Social Media Forum

17 May 2006, London Hilton Metropole

http://www.socialmediaforum.co.uk/

This one day forum will examine the impact of wikis, blogs & RSS on how organisations can monitor and control what is being said about them both internally and externally.

The programme includes:

- choosing the medium - wikis, blogs, RSS – where is the value?
- how you can communicate and collaborate in your organisation - case studies from Dresdner Kleinwort Wasserstein, Allen & Overy and Sun Microsystems
- the future of social media and web 2.0 - where's it all going.

The event will be chaired by Euan Semple, Independent Consultant and ex-Head of Knowledge Management at the BBC.

The Early Bird Delegate fee is £350 +VAT and includes all sessions, lunch, speaker notes plus access to the sponsor showcase and the blog and social media clinic. More information and online booking visit can be found on the website at http://www.socialmediaforum.co.uk/

Internet Librarian International 2006

16-17 October 2006, Copthorne Tara Hotel, London

http://www.internet-librarian.com/index.shtml

A reminder to you all that the deadline for submitting papers to Internet Librarian International in London is fast approaching. It may be taking place in October but submissions need to be in by March 31st. The organisers are looking for a mix of papers for conference sessions, workshops,

and short tutorials with the emphasis on the practical rather than theoretical. For example, particularly welcome will be case studies and proposals about initiatives in your library or information centre. A list of topics can be found on the Internet Librarian web site.

As someone who was involved with ILI last year, both as a speaker and a delegate, I can thoroughly recommend the event. I came away with lots of ideas as to how I can use new technologies such as blogs, RSS and wikis in my work and I was even encouraged to try Instant Messaging! (See my blog report on Brian Kelly's presentation "Email must die" - http://www.rba.co.uk/rss/2005/10/email-must-die.html)

European Firefox use hits 20 per cent

VNU.net's summary of a report from French internet monitoring firm XiTi Monitor (http://www.vnunet.com/vnunet/news/2148740/european-firefox-hits-per-cent) says that European Firefox usage has reached 20 per cent. Apparently "Finland leads the way with nearly 40 per cent of computer users choosing the open source browser, while the UK languishes near the bottom with just 11 per cent. Germany and Slovenia are the only other countries with over 30 per cent market share for the Firefox browser." My own web site stats (http://www.rba.co.uk/) now show that around 18% of visitors are Firefox fans and UKeiG - the UK eInformation Group reports that around 20% are using the alternative browser (http://www.ukeig.org.uk/blog/2006/02/ukeig-browser-share.html).

My RSS

http://www.my-rss.co.uk/

My RSS is a free web based service, set up by Steve Burgess, that enables you to quickly create your own RSS feeds. The interface is very easy to use and ensures that you fill in all the "required" fields. The resulting feed is hosted on the My RSS servers so you just point your users at the URL given by My RSS to your feed.

You can host up to 10 RSS feeds on the server and choose how to sort your feed and the number of items the feed has. In addition to being able to manually add items of news to your feed using the form, you can now "scrape" news from web sites. All you need to do is supply the URL of the news item and the website will extract all the information it can from the page. The site is standards compliant (XHTML and CSS) and the feeds produced by it are compliant with RSS 2. Feeds are also categorised using the DMOZ Open Directory Project classification scheme (http://dmoz.org/).

The great thing about this service is that you do not have to know much about creating RSS feeds. It does most of the work for you - you just have to type or paste the information into the relevant boxes.

Information Resources

Russia: All Regions, Trade & Investment Guide

http://www.dataresources.co.uk/

Available from Effective Technology Marketing, Russia: All Regions, Trade and Investment Guide is a comprehensive compilation of data for all regions of Russia. The Guide provides accurate, up-to-date information on markets, current economic conditions, sources of supply, infrastructure, trade opportunities, operating conditions, investment projects, legal and tax environment.

Data is gathered directly from local administrations and governments, and reviewed and verified through government agencies. The Guide has been completely revised for 2006 and is available separately in English and Russian with full colour charts, graphs and maps, and a fully searchable CD-ROM. There is a prepublication discount of 25% and discounted prices range from Eur 449 to Eur 679 depending on whether you opt for the book and/or CD-ROM and the language of publication (English or Russian, or both)

BvDEP launches Russian company database

http://www.bvdep.com/

Bureau van Dijk Electronic Publishing (BvDEP) have launched RUSLANA, a source of information on companies in Russia and the Ukraine. RUSLANA contains both standardised and "as reported" financial information for nearly a million companies. RUSLANA includes profit and loss and balance sheet data in various accounting formats for approximately 600,000 Russian, and over 260,000 Ukrainian, companies, with summaries for a further 100,000 Ukrainian companies.

The data is supplied by Creditreform, providers of business reports on companies in both regions. Supplementary data includes:

- BvDEP detailed ownership information
- news including M&A activity from ZEPHYR
- activities
- management
- import/export details
- stock price data

RUSLANA is provided via the internet initially in English but with plans to launch a Cyrillic version in the near future. It is aimed at financial services/banks, corporates, consultancies and accountants, as well as the public sector, and can be used for financial research, credit analysis and business to business marketing. More information about RUSLANA is available from BvDEP at info@bvdep.com and bvdep.com where you can request a free trial.

Accoona is back

http://www.accoona.com/

When Accoona was first released in December 2004 it failed to impress. Much was made of the artificial intelligence that is part of the system. However, search results were often anything but intelligent. This new version is greatly improved, although the business search is just a D&B search on US companies. It is the News section that I find particularly interesting with its "SuperTarget Your Search" option. Publication coverage is excellent and includes a good range of industry sectors and international, regional and local sources. My standard test search on wind farms for example picked up the Taunton Gazette and Rutland Herald as well as The Guardian, Washington Post and a host of environmental publications.

Results can be sorted by date and relevance but the SuperTarget goes even further and enables you to narrow down your search by publisher, companies and people mentioned in the articles, country and state. The "When published" feature includes an option for "More than 30 days ago" but this does limit you to only those sources that offer free access beyond this date.

This is a worthy competitor to the free offerings from Google, Yahoo and Ask but it currently lacks an Alerts option.

Searching Questions

Looking for free-to-use images

Question:

I am on the committee of a small local charity and am looking for pictures of a range of impressive looking buildings for a publicity brochure. I know the names and locations of the buildings I want to include but when I carry out image searches in the main search engines, I then have to try and track down the page owner to find out the copyright and license status of the photos. They sometimes come back to me and say that the image is copyright free because they found it on the web. Is that right? And if not, where can I be sure of finding images that are free to use for my purposes?

Answer:

Oh dear! They have indeed got it wrong. Just because an image appears on the web does not mean that anyone can copy and use it for any purpose, private or commercial. You do need to contact the owner of the page and the photo and ascertain how and where you can use it. There are, though, some sites that you can go to where the "rights" status is clearly stated for each photo and how it can be used.

Flickr - http://www.flickr.com/. Photo collections made available by individuals for sharing within closed groups or the world in general. Can be difficult to search unless you can work out what tags or keywords have been assigned to the images by the photographer. Go to http://www.flickr.com/photos/search/ and use the search box for titles, tags and descriptions. Each photo will have a copyright and licensing statement associated with it indicating what you can or cannot do with it. As a short cut, you can search for photos that have been assigned a specific Creative Commons license at http://flickr.com/creativecommons/ . Details of the different licenses are on that page as well.

Morguefile - http://www.morguefile.com/. No, this is not a collection of photos of bodes and corpses. Morgue file is a journalists term for a news archive. This site contains free high resolution digital stock photography for either corporate or public use.

Wikimedia Commons - http://commons.wikimedia.org/ . All the images, and other content, on this web site are free to use. Everyone is allowed to copy, use and modify any files as long as the source and the authors are credited, and as long as you release your copies/improvements under the same freedom to others.

These things are sent to try us

The case of the Google asterisk and the missing 500,000 pages

I came across an interesting glitch in Google this week recently. I often use the asterisk in a search strategy in Google to stand in for one or more words in a phrase. Originally an asterisk between two words in a phrase stood in for one term, two asterisks represented two terms etc. So, for example, "climate * change" would look for the words climate and change separated by one word, and "climate * * change" would find the two terms separated by two words. If the quote marks are removed then a single asterisk can represent one or more words.

It quickly became apparent, and was later confirmed by Google, that it does not matter whether there are quote marks around the phrase or not and that you need use only one asterisk to stand in for 0, 1 or more words. But that was not the problem that concerned me. One of my regular series of searches that I use in demonstrations and on training courses (phenol extraction) was coming up with some really weird results. The standard types of searches are OK, for example:

phenol extraction about 9,700,000
"phenol extraction" - about 141,000

Both of these are as expected and I can browse through several pages of results if I want to.

However:

phenol * extraction - would only display 1 - 2 of about 441,000
phenol * * extraction - displayed only 1 - 5 of about 513,000

I mentioned this in the AIIP (Association of Independent Information Professionals) and UKeiG discussion lists and another member came up with yet more variations using "Polish notation":

phenol extraction * 1-49 of about 207,000
phenol extraction * * 1-44 of about 207,000

Why had 200,000 - 500,000 of my results gone AWOL? I had no desire to look at them all, and in any case you can only view up to about a 1000 in Google, but I was curious. Google did respond to my report and confirmed that you use a single asterisk between words without quote marks to represent one or more words. Additionally you can use it after a series of words to fill in the blanks for a query that corresponds to a question. They also said that searches sometimes return "erroneous" estimates for the number of results for a search, which most of us knew already through bitter experience. They do not, though, explain why phenol * extraction still only displays 1-2 of 471,000!

Gary Price has since directed me to a posting on Greg Notess's site at
http://www.searchengineshowdown.com/features/google/inconsistent.shtml .The phenomenon has been known for a while and is called a GoogleNACK (Negative ACKnowledgements). In essence, it confirms that you cannot trust Google, and probably most of the other search engines, to return accurate results (What a surprise - NOT). So, I have now inserted a slide into my standard presentation on search engines that says:

"Do not attempt to apply logic to Google's search results - therein lies the path to madness".

Gizmo of the Month

FeedSpring

http://www.usablelabs.com/productFeedSpring.html

Need to create an RSS feed from scratch? FeedSpring is free feed generator software, simple to use and ensures that you include all the essential elements. Just simply type RSS information into required and optional fields in the forms it presents and FeedSpring automatically generate an RSS feed file for you. Then all you need to do is load the file into your web space, put a link to it on your web page and that's it.

FeedSpring currently supports only RSS 2.0. The current release runs on MS Windows but versions for Linux, FreeBSD, and MacOS will also be provided in the near future.

Meetings and Workshops

Workshop: Developing and managing e-book collections
Organiser: UKeiG
Presenter: Chris Armstrong, Ray Lonsdale
Venue: Netskills Training Suite, University of Newcastle Upon Tyne, Newcastle
Date: Tuesday, 11th April 2006, 9.30 – 16.30
Course fee: UKeiG members £150 + VAT (£176.25); others £180 + VAT (£211.50)
URL: http://www.ukeig.org.uk/training/2005_04_11_ebook.html
Outline: This course opens the door to a new electronic format. In the last six years, there has been an unprecedented growth in the publishing of e-books with an increasing array of different types available for all sectors. The programme will give you the opportunity to explore a range of different e-books including a range of commercially-published and free reference works, monographs, textbooks, and fiction. Examples will include individual titles and also collections of e-books, such as those offered by NetLibrary and Oxford University Press. The course will also facilitate consideration of the new opportunities e-books offer for librarians and users, and the significant collection management and promotional issues which challenge information and library staff.

Workshop: Electronic Information Risk Management
Organiser: UKeiG
Presenter: Dr Rita Esen
Venue: CILIP, 7 Ridgmount Street, London, WC1E 7AE
Date: Tuesday 16th May 2006, 10.00 – 16.30
Course fee: UKeiG members £130 + VAT (£152.75); others £160 + VAT (£188.00).
URL: http://www.ukeig.org.uk/training/may06/einfo_riskman.html
Outline: Today's networked economy has changed the way organisations operate as they are now faced with the challenges of using electronic information to fulfil their business goals. Electronic Information Risk Management consists of the use of best practice to collect, organise, use, store, share, provide access to and dispose of electronic information. Every organisation is now required to ensure that the use of electronic information and e-systems comply with legal, regulatory and best practice requirements. This training course will provide a sound understanding of electronic information risks and how to manage them. It will be a combination of presentations, group tasks, discussions of best practice and practical problem solving sessions. A practical case study will be used to highlight typical areas of e-information risks.

Workshop: Making Websites Accessible
Organiser: UKeiG
Presenter: Nigel Ford, Peter Holdridge
Venue: Department of Information Studies, University of Sheffield
Date: Wednesday, 24th May 2006, 9.45 - 16.30
Course fee: UKeiG members £150 + VAT (£176.25); others £180 + VAT (£211.50)
URL: http://www.ukeig.org.uk/training/2006_05_24_accessiblesites.html
Outline: It is now a legal requirement to ensure that Websites are designed in such a way as to make them accessible to people with a range of disabilities. In the UK, the Disability Discrimination Act (DDA), stipulates that websites must be accessible to blind and disabled users. This is a one-day "hands on" practical course designed to teach you how to create Web pages and Websites that are accessible to people with a range of disabilities. You will learn how to validate pages that you create in relation to acknowledged international standards, and how to kite mark your Web pages according to A, AA or AAA international accessibility standards. You will become familiar with techniques and software that provides automatic checking of your Web pages for accessibility, and will learn how to keep up to date with future accessibility developments.

Contact Information

Karen Blakeman, RBA Information Services
UK Tel: 0118 947 2256
Int. Tel: +44 118 947 2256

UK Fax: 020 8020 0253
Int. Fax: +44 20 8020 0253

Address: 88 Star Road, Caversham, Berks RG4 5BE, UK

Archives

TFTTR archives: http://www.rba.co.uk/tfttr/archives/

Subscribe and Unsubscribe

To subscribe to the newsletter fill in the online registration form at http://www.rba.co.uk/tfttr/

To unsubscribe, use the registration form at http://www.rba.co.uk/tfttr/ and check the unsubscribe radio button.

If you wish to change your subscription from the PDF to the Plain Text version, send an email to tfttr@rba.co.uk with "Plain Text Version Please" in the subject line.

Privacy Statement

Subscribers' details are used only to enable distribution of the newsletter **Tales from the Terminal Room**. The subscriber list is not used for any other purpose, nor will it be disclosed by RBA Information Services or made available in any form to any other individual, organisation or company.